Inverse problem robustness for multi-species mean field spin models

Micaela Fedele[1], Cecilia Vernia[2], and Pierluigi Contucci[3]

[1] *Courant Institute of Mathematical Sciences, New York University*

[2]*Dipartimento di Scienze Fisiche Informatiche e Matematiche, Università di Modena e Reggio Emilia*

[3]*Dipartimento di Matematica, Università di Bologna*

Abstract

The inverse problem method is tested for a class of mean field statistical mechanics models representing a mixture of particles of different species. The robustness of the inversion is investigated for different values of the physical parameters, system sizes and independent samples. We show how to reconstruct the parameter values with a precision of a few percentages.

Keywords: *inverse problem in statistical mechanics, multi-species mean field model, maximum likelihood procedure, finite size effects, Curie-Weiss models*

1 Introduction

In recent years there has been an increasing interest in studying the inverse problem in statistical mechanics mostly due to the fact that the thermodynamic formalism on a macroscopic base has proved to be effective in a variety of scientific applications that span

from the investigation of real neural networks in biology [17–19] to behavioral ethology for flocks [1]. In this paper we are interested in a particular class of models [3,8] that have naturally emerged within the application of the statistical mechanics formalism to socio-economic sciences (see also [16,20] and references therein). Their first and most elementary appearance can be traced back to the so called discrete choice theory proposed by Daniel Mc Fadden [15] after his celebrated success in predicting the number and distribution of customers of the Bay Area Rapid Transit before its construction. Discrete choice theory doesn't contain interaction between individuals and from the statistical mechanics point of view can be seen as a mixture of a finite number of discrete perfect gases; its inverse problem is mathematically elementary and its efficiency amounts to the proper identification of the different species of particles (see also [11]). The necessity to include the interaction among agents led W. Brock and S.N. Durlauf [2] to introduce, within the socio-economic context, the simplest interaction structure which is given by the mean field Curie-Weiss Hamiltonian model. In order to successfully generalize the discrete choice to the interacting case, it was defined in [3] a multi-species mean field model. In this paper we propose a robustness test of the inverse problem in the multi-species mean field case. We start from the knowledge of the exact solution of the model, both in the single populated system and in the bi-populated one, not only at the thermodynamic limit in analytic form but also at finite and increasing number of particles by accurate numerical approximations. This, together with the standard criterion of maximum likelihood, provides a relation between experimental and theoretical quantities and allows to tackle the computation of the free parameters of the model, namely interactions and magnetic fields, from observed data.

We then generate the equilibrium configuration of the models, at different system sizes and for different values of the parameters i.e. interaction strength and magnetic field. By use of the inversion formulas we show how the reconstruction of the parameters is achieved and how his robustness depends on both system size and number of independent samples used.

The paper is organized as follows. In section 2 we recall briefly the Curie-Weiss model

and we review how to solve the inverse problem in this single-population model. The generalization of such a model to systems composed of many interacting groups (the multi-species mean field model) and the solution of the corresponding inverse problem are presented in section 3. Section 4 presents and discusses a set of numerical tests for both the single and the bi-populated system for finite number of particles and finite number of samples. We first investigate how the average quantities, magnetization and susceptibility behave for increasing system sizes in the standard Curie Weiss model. We find, in particular, that while the magnetization is monotonically increasing, in agreement with the first Griffiths Kelly Sherman inequality [9, 10, 12], the susceptibility has a monotonicity direction that changes with the values of the coupling constant with respect to its critical point. Both quantities reach their limiting value at the speed of the inverse volume. We then investigate how the experimental magnetization and susceptibility at fixed volume depend on the number of samples and stabilize when their number increases. The effectiveness of the inversion is tested for different values of the coupling constants and magnetic fields. The same procedure is applied to the bi-populated model and again the robustness of the inversion is tested for different values of the parameters. We find in all cases that the inverse method reconstructs, with a modest amount of samples, the values of the parameters with a precision of a few percentages.

2 Inverse problem for the Curie-Weiss model

Denoting with N the size of the population, the Curie-Weiss model is defined by the Hamiltonian:

$$H_N(\boldsymbol{\sigma}) = -\frac{J}{2N} \sum_{i,j=1}^{N} \sigma_i \sigma_j - h \sum_{i=1}^{N} \sigma_i \tag{1}$$

where $\sigma_i \in \{\pm 1\}$ is the spin of the particle i (individual), the parameter $J > 0$ is the coupling constant and h is the value of the magnetic field. The joint probability of a

configuration of spins $\boldsymbol{\sigma} = (\sigma_1, \ldots, \sigma_N)$ is given by the Boltzmann-Gibbs measure:

$$P_{N,J,h}\{\boldsymbol{\sigma}\} = \frac{\exp(-H_N(\boldsymbol{\sigma}))}{\sum\limits_{\boldsymbol{\sigma} \in \Omega_N} \exp(-H_N(\boldsymbol{\sigma}))}$$

$$= \frac{\exp\left(N\left(\frac{J}{2}m_N^2(\boldsymbol{\sigma}) + hm_N(\boldsymbol{\sigma})\right)\right)}{\sum\limits_{\boldsymbol{\sigma} \in \Omega_N} \exp\left(N\left(\frac{J}{2}m_N^2(\boldsymbol{\sigma}) + hm_N(\boldsymbol{\sigma})\right)\right)} \tag{2}$$

where $\Omega_N = \{-1, 1\}^N$ and

$$m_N(\boldsymbol{\sigma}) = \frac{1}{N}\sum_{i=1}^{N} \sigma_i$$

is the magnetization of the configuration $\boldsymbol{\sigma}$. We point out that, in the inverse problem, the usual inverse temperature parameter β is absorbed within the two free parameters J and h.

We will denote by $\omega(\cdot)$ the expectation value with respect to $P_{N,J,h}$. Heuristically, this distribution favors both the agreement of people's choices or opinions σ_i, with an external influence h, and the agreement between individuals, being J positive (whereas for $J < 0$ it would favor disagreement).

The inverse problem amounts to compute the values of J and h starting from the knowledge of the magnetization average and variance. Of course, when dealing with real phenomenological data its solution is made in two steps. The first is the identification of the analytical inverse formula providing a possible explicit expression of the free parameters (J and h) in terms of the mentioned macroscopic thermodynamic variables. The second is the evaluation with statistical methods of the macroscopic variables starting from real data. The problem is generically well posed because the unknown parameters (interaction and magnetic field) are as many as the measured phenomenological quantities (average magnetization and its fluctuation).

Let start by observing that when $h \neq 0$ and $J > 0$ or $h = 0$ and $J < 1$, the Curie-Weiss model satisfies the following property (see [4])

$$\lim_{N \to \infty} \omega(m_N(\boldsymbol{\sigma})) = m(J, h) \tag{3}$$

4

where $m(J, h)$ is the stable solution of the model mean-field equation

$$m(J, h) = \tanh(Jm(J, h) + h). \tag{4}$$

By differentiating the identity (3) with respect to h we obtain:

$$\lim_{N \to \infty} \frac{\partial}{\partial h} \omega(m_N(\boldsymbol{\sigma})) = \chi$$

where

$$\chi = \frac{\partial m(J, h)}{\partial h} = \frac{1 - m^2(J, h)}{1 - J(1 - m^2(J, h))} \tag{5}$$

and

$$\frac{\partial}{\partial h} \omega(m_N(\boldsymbol{\sigma})) = \frac{\partial}{\partial h} \left(\frac{\sum_{\boldsymbol{\sigma} \in \Omega_N} m_N(\boldsymbol{\sigma}) \exp(-H_N(\boldsymbol{\sigma}))}{\sum_{\boldsymbol{\sigma} \in \Omega_N} \exp(-H_N(\boldsymbol{\sigma}))} \right)$$
$$= N\Big(\omega(m_N^2(\boldsymbol{\sigma})) - \omega^2(m_N(\boldsymbol{\sigma}))\Big). \tag{6}$$

In particular, the right hand side of the last identity defines the finite size susceptibility χ_N. By putting together (5) and (6) we can compute the parameter J from the average value and the variance of the magnetization in the thermodynamic limit:

$$J = \frac{1}{1 - \lim_{N \to \infty} \omega^2(m_N(\boldsymbol{\sigma}))} - \frac{1}{\lim_{N \to \infty} N\Big(\omega(m_N^2(\boldsymbol{\sigma})) - \omega^2(m_N(\boldsymbol{\sigma}))\Big)}. \tag{7}$$

The external field h is obtained, in the large volume limit, by inverting the mean-field equation (4)

$$h = \tanh^{-1}\Big(\lim_{N \to \infty} \omega(m_N(\boldsymbol{\sigma}))\Big) - J \lim_{N \to \infty} \omega(m_N(\boldsymbol{\sigma})) \tag{8}$$

where J is given by (7). Formulas (7) and (8) solve the analytical inverse problem for the Curie-Weiss model as $h \neq 0$ and $J > 0$ or $h = 0$ and $J < 1$. On the other hand, if $h = 0$ and $J > 1$ the equation (4) has two different stable solutions $\pm m(J, 0)$ (see [4]). Thus, the identity (3) is not verified. In this case the inverse problem can be solved, by observing that there exists $\epsilon > 0$ such that, whenever $m_N(\boldsymbol{\sigma}) \in (\pm m(J, 0) - \epsilon, \pm m(J, 0) + \epsilon)$, the following limit holds (see [5])

$$\lim_{N \to \infty} \omega(m_N(\boldsymbol{\sigma})) = \pm m(J, 0) \tag{9}$$

5

and then by applying to (9) the same procedure as shown above. The result is still given by formulas (7) and (8) which conclude the analytical treatment of the inverse problem.

For what it concerns the statistical part one has to provide an evaluation of the finite size average magnetization $\omega(m_N(\boldsymbol{\sigma}))$ and susceptibility χ_N from the empirical data. We use the maximum likelihood estimation procedure. This method identifies the free parameters within a distribution by requiring that their value maximize the probability to obtain the given sample, under the condition that the sample is made of independent and identically distributed realizations of the random variables.

Given a sample made of M independent spin configurations $\boldsymbol{\sigma}^{(1)}, \ldots, \boldsymbol{\sigma}^{(M)}$ all distributed according to the Boltzmann-Gibbs measure (2), the maximum likelihood [7, 13] function is defined by

$$L(J,h) = P_{N,J,h}\Big\{\boldsymbol{\sigma}^{(1)}, \ldots, \boldsymbol{\sigma}^{(M)}\Big\}$$

which, using the independence, can be rewritten as

$$L(J,h) = \prod_{i=1}^{M} P_{N,J,h}\Big\{\boldsymbol{\sigma}^{(i)}\Big\}$$

$$= \prod_{i=1}^{M} \frac{\exp(-H_N(\boldsymbol{\sigma}^{(i)}))}{\sum_{\boldsymbol{\sigma} \in \Omega_N} \exp(-H_N(\boldsymbol{\sigma}))}.$$

To maximize the function $L(J,h)$ we should compute the derivative of a product. Since a function and its logarithm reach the maximum in the same point, we consider the logarithm of the maximum likelihood function

$$\ln L(J,h) = \sum_{i=1}^{M} \Big(-H_N(\boldsymbol{\sigma}^{(i)}) - \ln \sum_{\boldsymbol{\sigma} \in \Omega_N} \exp(-H_N(\boldsymbol{\sigma})) \Big).$$

The derivatives with respect to h and J of this function

$$\frac{\partial \ln L(J,h)}{\partial h} = N \sum_{i=1}^{M} \Big(m_N(\boldsymbol{\sigma}^{(i)}) - \omega(m_N(\boldsymbol{\sigma})) \Big)$$

$$\frac{\partial \ln L(J,h)}{\partial J} = \frac{N}{2} \sum_{i=1}^{M} \Big(m_N^2(\boldsymbol{\sigma}^{(i)}) - \omega(m_N^2(\boldsymbol{\sigma})) \Big)$$

vanish for

$$\begin{cases} \omega(m_N(\boldsymbol{\sigma})) = \dfrac{1}{M} \sum_{i=1}^{M} m_N(\boldsymbol{\sigma}^{(i)}) \\[4mm] \omega(m_N^2(\boldsymbol{\sigma})) = \dfrac{1}{M} \sum_{i=1}^{M} m_N^2(\boldsymbol{\sigma}^{(i)}). \end{cases} \tag{10}$$

Therefore, the function $L(J, h)$ reaches its maximum when the first and second momentum of the magnetization are calculated from the data according to (10). The inverse problem is finally solved by the composition of (10) with (7) and (8). In particular, denoting by m_{exp} and χ_{exp} respectively the average magnetization and the susceptibility computed from the sample

$$m_{exp} = \frac{1}{M} \sum_{i=1}^{M} m_N(\boldsymbol{\sigma}^{(i)}) \qquad \chi_{exp} = N \left(\frac{1}{M} \sum_{i=1}^{M} m_N^2(\boldsymbol{\sigma}^{(i)}) - m_{exp}^2 \right), \tag{11}$$

the estimators of the model's free parameters are

$$J_{exp} = \frac{1}{1 - m_{exp}^2} - \frac{1}{\chi_{exp}} \tag{12}$$

$$h_{exp} = \tanh^{-1}(m_{exp}) - J_{exp} m_{exp}. \tag{13}$$

3 Inverse problem for the multi-species model

Since, within the applications we are interested in, we aim at generalizing the discrete choice theory [15], we proceed toward the solution of the inverse problem for the multi-species mean field model. Formally such a model is an extension of the Curie-Weiss model to systems composed of many interacting groups. We consider a system of N particles that can be divided into k subsets $P_1, \ldots, P_k$ with $P_l \cap P_s = \emptyset$, for $l \neq s$ and sizes $|P_l| = N_l$, where $\sum_{l=1}^{k} N_l = N$. Particles interact with each other and with an external field according to the mean field Hamiltonian:

$$H_N(\boldsymbol{\sigma}) = -\frac{1}{2N} \sum_{i,j=1}^{N} J_{ij} \sigma_i \sigma_j - \sum_{i=1}^{N} h_i \sigma_i . \tag{14}$$

The $\sigma_i \in \{\pm 1\}$ represents the spin of the particle i, while J_{ij} is the parameter that tunes the mutual interaction between the particle i and the particle j and takes values according to the following symmetric matrix:

$$
\begin{array}{c}
N_1 \left\{ \vphantom{\begin{array}{c} a \\ b \end{array}} \right. \\
N_2 \left\{ \vphantom{\begin{array}{c} a \\ b \end{array}} \right. \\
\\
N_k \left\{ \vphantom{\begin{array}{c} a \\ b \\ c \end{array}} \right.
\end{array}
\overbrace{\qquad}^{N_1} \overbrace{\qquad}^{N_2} \quad \overbrace{\qquad\qquad}^{N_k}
\left(
\begin{array}{c c | c | c}
\mathbf{J}_{11} & \mathbf{J}_{12} & \cdots & \mathbf{J}_{1k} \\ \hline
\mathbf{J}_{12} & \mathbf{J}_{22} & & \\ \hline
\vdots & & & \\ \hline
\mathbf{J}_{1k} & \mathbf{J}_{2k} & \cdots & \mathbf{J}_{kk}
\end{array}
\right)
$$

where each block $\mathbf{J}_{ls}$ has constant elements J_{ls}. For $l = s$, $\mathbf{J}_{ll}$ is a square matrix, whereas the matrix $\mathbf{J}_{ls}$ is rectangular. We assume $J_{11}, J_{22}, \ldots, J_{kk}$ to be positive, whereas J_{ls} with $l \neq s$ can be either positive or negative allowing for both ferromagnetic and antiferromagnetic interactions. The vector field takes also different values depending on the subset the particles belong to as specified by the following vector:

$$
\begin{array}{c}
N_1 \left\{ \vphantom{a} \right. \\
N_2 \left\{ \vphantom{a} \right. \\
\\
N_k \left\{ \vphantom{a} \right.
\end{array}
\left(
\begin{array}{c}
\mathbf{h}_1 \\ \hline
\mathbf{h}_2 \\ \hline
\vdots \\ \hline
\mathbf{h}_k
\end{array}
\right)
$$

where each $\mathbf{h}_l$ is a vector of constant elements h_l. Indicating with $m_l(\boldsymbol{\sigma})$ the magnetization of the group P_l, and with $\alpha_l = N_l/N$ the relative size of the set P_l, we may easily express

the Hamiltonian (14) as:

$$H_N(\boldsymbol{\sigma}) = -N\left(\frac{1}{2}\sum_{l,s=1}^{k} \alpha_l \alpha_s J_{ls} m_l(\boldsymbol{\sigma}) m_s(\boldsymbol{\sigma}) + \sum_{l=1}^{k} \alpha_l h_l m_l(\boldsymbol{\sigma})\right)$$

$$= -N\left(\frac{1}{2}\langle \mathbf{J}\mathbf{D}_{\boldsymbol{\alpha}}\mathbf{m}(\boldsymbol{\sigma}), \mathbf{D}_{\boldsymbol{\alpha}}\mathbf{m}(\boldsymbol{\sigma})\rangle + \langle \mathbf{h}, \mathbf{D}_{\boldsymbol{\alpha}}\mathbf{m}(\boldsymbol{\sigma})\rangle\right) \tag{15}$$

where $\mathbf{m}(\boldsymbol{\sigma}) = (m_1(\boldsymbol{\sigma}), \ldots, m_k(\boldsymbol{\sigma}))$, $\mathbf{D}_{\boldsymbol{\alpha}} = diag\{\alpha_1, \ldots, \alpha_k\}$, $\mathbf{h} = (h_1, \ldots, h_k)$ and $\mathbf{J}$ is the reduced interaction matrix

$$\mathbf{J} = \begin{pmatrix} J_{11} & J_{12} & \ldots & J_{1k} \\ J_{12} & J_{22} & \ldots & J_{2k} \\ \vdots & \vdots & & \vdots \\ J_{1k} & J_{2k} & \ldots & J_{kk} \end{pmatrix}.$$

The joint distribution of a spin configuration $\boldsymbol{\sigma} = (\sigma_1, \ldots, \sigma_N)$ is given by the Boltzmann-Gibbs measure $P_{N,\mathbf{J},\mathbf{h}}$ related to the Hamiltonian (14), where again we consider the inverse temperature parameter β absorbed within the model parameters $\mathbf{J}$ and $\mathbf{h}$. The well position of the model has been shown in [8]. In particular, in the thermodynamic limit the model is described by the following system of mean-field equations:

$$\begin{cases} m_1(\mathbf{J}, \mathbf{h}) = \tanh\left(\sum_{l=1}^{k} \alpha_l J_{1l}\, m_l(\mathbf{J}, \mathbf{h}) + h_1\right) \\ m_2(\mathbf{J}, \mathbf{h}) = \tanh\left(\sum_{l=1}^{k} \alpha_l J_{2l}\, m_l(\mathbf{J}, \mathbf{h}) + h_2\right) \\ \vdots \\ m_k(\mathbf{J}, \mathbf{h}) = \tanh\left(\sum_{l=1}^{k} \alpha_l J_{lk}\, m_l(\mathbf{J}, \mathbf{h}) + h_k\right). \end{cases} \tag{16}$$

If the system (16) admits a unique thermodynamically stable solution $\mathbf{m}(\mathbf{J}, \mathbf{h}) = (m_1(\mathbf{J}, \mathbf{h}), \ldots, m_k(\mathbf{J}, \mathbf{h}))$, the following identities hold (see [6]):

$$\lim_{N\to\infty} \omega(m_l(\boldsymbol{\sigma})) = m_l(\mathbf{J}, \mathbf{h}) \quad l = 1, \ldots, k. \tag{17}$$

By differentiating the identities (17) with respect to h_s, $s = 1, \ldots, k$, we obtain

$$\lim_{N\to\infty} \frac{\partial}{\partial h_s}\omega(m_l(\boldsymbol{\sigma})) = \chi_{ls} \quad l, s = 1, \ldots, k \tag{18}$$

where χ_{ls} are the elements of the susceptibility matrix of the model. In particular,

$$\chi_{ls} = \frac{\partial m_l(\mathbf{J}, \mathbf{h})}{\partial h_s} = \frac{\partial}{\partial h_s}\left(\tanh\left(h_l + \sum_{p=1}^{k} \alpha_p J_{lp} m_p(\mathbf{J}, \mathbf{h})\right)\right)$$

$$= (1 - m_l^2(\mathbf{J}, \mathbf{h}))\left(\delta_{ls} + \sum_{p=1}^{k} \alpha_p J_{lp} \chi_{ps}\right)$$

where δ_{ls} denotes the delta of Dirac picked in $l = s$. Therefore, the susceptibility matrix $\boldsymbol{\chi}$ can be written as:

$$\boldsymbol{\chi} = \mathbf{P}(\mathbf{I} + \mathbf{J}\mathbf{D}_{\boldsymbol{\alpha}}\boldsymbol{\chi}) \tag{19}$$

where $\mathbf{P} = diag\{1 - m_1^2(\mathbf{J}, \mathbf{h}), \ldots, 1 - m_k^2(\mathbf{J}, \mathbf{h})\}$ and $\mathbf{I}$ is the identity matrix. Moreover, for each $l, s = 1, \ldots, k$

$$\frac{\partial}{\partial h_s}\omega(m_l(\boldsymbol{\sigma})) = \frac{\partial}{\partial h_s}\left(\frac{\sum_{\boldsymbol{\sigma} \in \Omega_N} m_l(\boldsymbol{\sigma})e^{-H_N(\boldsymbol{\sigma})}}{\sum_{\boldsymbol{\sigma} \in \Omega_N} e^{-H_N(\boldsymbol{\sigma})}}\right)$$

$$= N_s\Big(\omega(m_l(\boldsymbol{\sigma})m_s(\boldsymbol{\sigma})) - \omega(m_l(\boldsymbol{\sigma}))\omega(m_s(\boldsymbol{\sigma}))\Big). \tag{20}$$

By computing the elements of $\boldsymbol{\chi}$ according to (18) and (20), by (19) we get an expression of the reduced interaction matrix $\mathbf{J}$ related to the average value and the correlations of the magnetizations in the thermodynamic limit:

$$\mathbf{J} = (\mathbf{P}^{-1} - \boldsymbol{\chi}^{-1})\mathbf{D}_{\boldsymbol{\alpha}}^{-1} , \tag{21}$$

see [14]. Once the matrix $\mathbf{J}$ is determined, the elements of the vector $\mathbf{h} = (h_1, \ldots, h_k)$ are obtained by inverting the mean field equations (16)

$$h_l = \tanh^{-1}\left(\lim_{N\to\infty}\omega(m_l(\boldsymbol{\sigma}))\right) - \sum_{s=1}^{k} \alpha_s J_{ls} \lim_{N\to\infty}\omega(m_s(\boldsymbol{\sigma})) \quad l = 1, \ldots, k. \tag{22}$$

The previous formulas (21) and (22) represent the analytical solution of the inverse problem when the system of mean-field equations (16) has a unique stable solution. When there are more stable solutions, identities (17) have to be handled with care. Similar identities are, in fact, locally fulfilled around each solution and the solution of the inverse problem is still possible by applying to them the same procedure described above. The estimators are again given by (21) and (22). Before proceeding with the statistical part

10

of the inverse problem we notice that it is well posed since, in particular, the degrees of freedom of the problem are equal to $k(k+3)/2$.

Also in this case, starting from phenomenological data, we proceed with the help of the maximum likelihood principle. Consider a sample of M independent spin configurations $\boldsymbol{\sigma}^{(1)}, \ldots, \boldsymbol{\sigma}^{(M)}$ distributed according to the Boltzmann-Gibbs measure $P_{N,\mathbf{J},\mathbf{h}}$; the maximum likelihood function related to the sample is

$$
\begin{aligned}
L(\mathbf{J},\mathbf{h}) &= P_{N,\mathbf{J},\mathbf{h}}\Big\{\boldsymbol{\sigma}^{(1)}, \ldots, \boldsymbol{\sigma}^{(M)}\Big\} \\
&= \prod_{i=1}^{M} P_{N,\mathbf{J},\mathbf{h}}\Big\{\boldsymbol{\sigma}^{(i)}\Big\} \\
&= \prod_{i=1}^{M} \frac{\exp(-H_N(\boldsymbol{\sigma}^{(i)}))}{\sum_{\boldsymbol{\sigma}\in\Omega_N} \exp(-H_N(\boldsymbol{\sigma}))}.
\end{aligned}
\tag{23}
$$

Differentiating the logarithm of the likelihood function (23)

$$
\ln L(\mathbf{J},\mathbf{h}) = \sum_{i=1}^{M} \left(-H_N(\boldsymbol{\sigma}^{(i)}) - \ln \sum_{\boldsymbol{\sigma}\in\Omega_N} \exp(-H_N(\boldsymbol{\sigma})) \right)
$$

with respect to h_l and J_{ls}, $l,s = 1, \ldots, k$ we obtain

$$
\frac{\partial \ln L(\mathbf{J},\mathbf{h})}{\partial h_l} = N_l \sum_{i=1}^{M} \left(m_l(\boldsymbol{\sigma}^{(i)}) - \omega(m_l(\boldsymbol{\sigma})) \right)
$$

$$
\frac{\partial \ln L(\mathbf{J},\mathbf{h})}{\partial J_{ls}} = \frac{N\alpha_l\alpha_s}{2} \sum_{i=1}^{M} \left(m_l(\boldsymbol{\sigma}^{(i)})m_s(\boldsymbol{\sigma}^{(i)}) - \omega(m_l(\boldsymbol{\sigma})m_s(\boldsymbol{\sigma})) \right).
$$

These derivatives are equal to zero as the following equalities hold

$$
\begin{cases}
\omega(m_l(\boldsymbol{\sigma})) = \dfrac{1}{M} \sum\limits_{i=1}^{M} m_l(\boldsymbol{\sigma}^{(i)}) \quad l = 1, \ldots, k \\[4mm]
\omega(m_l(\boldsymbol{\sigma})m_s(\boldsymbol{\sigma})) = \dfrac{1}{M} \sum\limits_{i=1}^{M} m_l(\boldsymbol{\sigma}^{(i)})m_s(\boldsymbol{\sigma}^{(i)}) \quad l,s = 1, \ldots, k.
\end{cases}
\tag{24}
$$

Therefore, the inverse problem for the multi-species model is solved by the composition of (24) with (21) and (22). In particular, denoting by $m_{l\,exp}$ the average magnetization of each specie calculated from the data

$$
m_{l\,exp} = \frac{1}{M} \sum_{i=1}^{M} m_l(\boldsymbol{\sigma}^{(i)}) \quad l = 1, \ldots, k
$$

and defined the matrices $\mathbf{P}_{exp} = diag\{1 - m^2_{1\ exp}, \ldots, 1 - m^2_{k\ exp}\}$ and $\boldsymbol{\chi}_{exp}$, whose elements are

$$\chi_{ls\ exp} = N_s \left(\frac{1}{M} \sum_{i=1}^{M} m_l(\boldsymbol{\sigma}^{(i)}) m_s(\boldsymbol{\sigma}^{(i)}) - m_{l\ exp} m_{s\ exp} \right) \quad l, s = 1, \ldots, k$$

the model estimators are

$$\mathbf{J}_{exp} = (\mathbf{P}_{exp}^{-1} - \boldsymbol{\chi}_{exp}^{-1}) \mathbf{D}_{\boldsymbol{\alpha}}^{-1} \tag{25}$$

$$h_{l\ exp} = \tanh^{-1}(m_{l\ exp}) - \sum_{s=1}^{k} \alpha_s J_{ls\ exp} m_{s\ exp} \quad l = 1, \ldots, k. \tag{26}$$

4 The inversion at finite volume and finite sample size

When dealing with real data the elegant exactly solvable model has to be replaced by its finite size version. This is reflected both in the number of particles N and in the number M of independent configurations in the sample, available from the statistical set. It is therefore important to see how the inversion formulas perform for different values of those quantities at assigned values of the parameters. The Curie Weiss model and its generalized multi-species version provide an ideal testing set not only because most of the applications concern mean field models but also because their finite size solution can still be handled thanks to the observation that the magnetization spectrum has a probability distribution that can be exactly computed.

In this section we present a numerical test of our inversion procedure, both for the Curie-Weiss model and for its multi-species version. In both cases, for each choice of the size N of system and of the free parameter values (J and h for the Curie-Weiss model, $\mathbf{J}$ and $\mathbf{h}$ for its multi-species generalization), the data that we are going to use are extracted from a virtually exact simulation of the equilibrium distribution. This is possible thanks to the mean-field nature of the models (1) and (15), which reduces the computation of corresponding equilibrium distribution to that of the weights of the $O(N)$ values of the magnetization. In this way, from $P_{N,J,h}$ for the Curie-Weiss model and from $P_{N,\mathbf{J},\mathbf{h}}$ for

its multi-species generalization, we can compute the finite size average magnetization and susceptibility and extract sequences of configurations.

Although obvious, it is probably worth remarking that the parameter estimation method involves two approximations. The first one is in the inversion formulas (7),(8), (21) and (22), that require the infinite volume limit; the second one is the statistical error appearing in the evaluation of the averages and correlations through the maximum-likelihood estimators defined in (10) and (24). In principle, the first approximation could reduce strongly the scope of the method to systems with very large number N of individuals and it corresponds to estimating the finite size corrections of average magnetization and susceptibility. We don't go through this issue, rather we illustrate it with some numerical example to support the choice of the values of the parameters. Indeed, figure 1 shows the finite size average magnetization $\omega(m_N(\boldsymbol{\sigma}))$ and susceptibility $N(\omega(m_N^2(\boldsymbol{\sigma}))-\omega^2(m_N(\boldsymbol{\sigma})))$ for the Curie-Weiss model at different N's for $J = 0.6$, $h = 0.1$ and for $J = 1.2$, $h = 0.3$, while the same quantities in the thermodynamic limit $m(J, h)$ and χ are represented by the horizontal lines (to ease the notation, in the figure and in the following we denote the finite size quantities respectively by m_N and χ_N and we omit the dependences of the equilibrium magnetization). The figure highlights the monotonic behavior of m_N and χ_N as function of N. In particular, m_N is monotonic increasing for each value of the interacting parameter J, while χ_N is monotonic increasing as $J < 1$ and monotonic decreasing as $J > 1$. We point out that the different behavior of the finite size susceptibility is very useful dealing with empirical data because it tells us if the system is above or under the interacting parameter critical value before to apply the inversion procedure. Note that, for $N \geq 5000$ we have optimal approximations both for m and χ in the thermodynamic limit. The power-law fits in figure 2 show evidence of the $O(N^{-1})$-behavior of the finite size corrections both for magnetization and susceptibility, which entails the same $O(N^{-1})$ error in the estimation of J and h. Figure 3 gives evidence of the dependence of the estimators m_{exp} and χ_{exp}, given by (11), on the choice of the number M of the configurations of the sample that we use in the maximum-likelihood procedure. To asses the statistical dependence of m_{exp} and χ_{exp} on the sample $\{\boldsymbol{\sigma}^{(1)}, \ldots, \boldsymbol{\sigma}^{(M)}\}$, we computed
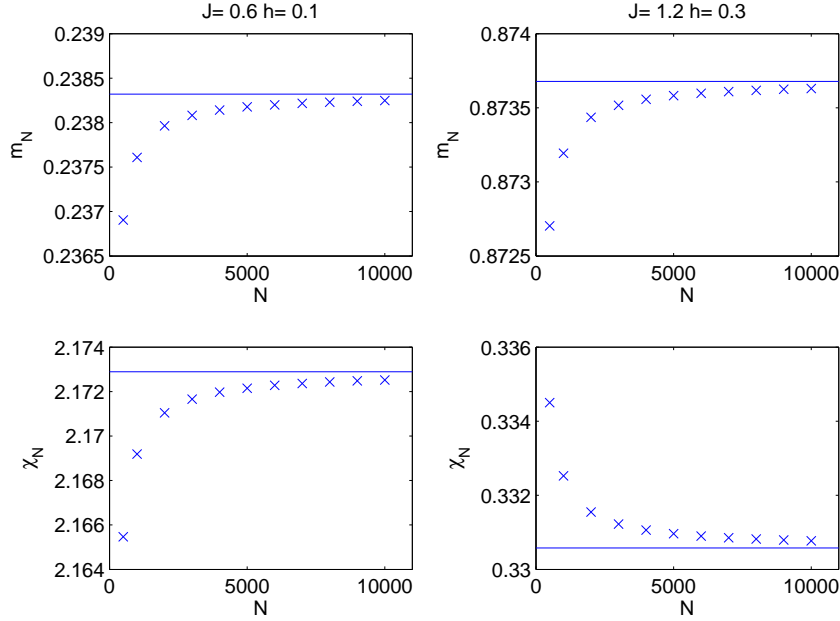
Figure 1: Finite size average magnetization m_N (upper panels) and susceptibility χ_N (lower panels) as a function of N for the Curie-Weiss model for $J = 0.6$ and $h = 0.1$ (left panels) and for $J = 1.2$ and $h = 0.3$ (right panel). The blue continuous lines represent the magnetization m and the susceptibility χ in the thermodynamic limit.

their values over a set of 20 independent instances of such samples. Thus from now on, and without ambiguity, we use the subscript exp to denote both estimators and their statistical mean over the 20 M-sample. We find numerical evidence that $M \geq 10000$ stabilizes the estimations. In particular, the standard deviation of both m_{exp} and χ_{exp} as a function of M behaves as a power law $M^{-0.5}$ (as $J = 0.6$ and $h = 0.1$, the fit of the standard deviation of m_{exp} is $aM^{-\alpha}$ with $\alpha = 0.4933 \pm 0.06$, $a = 0.013 \pm 0.006$, and goodness of fit $R^2 = 0.9696$, while those of the standard deviation of χ_{exp} is $bM^{-\beta}$ with $\beta = 0.5175 \pm 0.098$, $b = 3.269 \in (0.7388, 5.8)$ and $R^2 = 0.9343$).

In order to test numerically the inversion procedure for the Curie-Weiss model, we consider a sample of $M = 20000$ spin configurations $\{\boldsymbol{\sigma}^{(i)}\}, i = 1, \ldots, M$, where $\boldsymbol{\sigma}^{(i)} = (\sigma_1^{(i)}, \ldots, \sigma_N^{(i)})$ and $N = 10000$. For a given couple of parameters (J, h), we extract the sample of M independent identically distributed spin configurations from the Boltzmann-
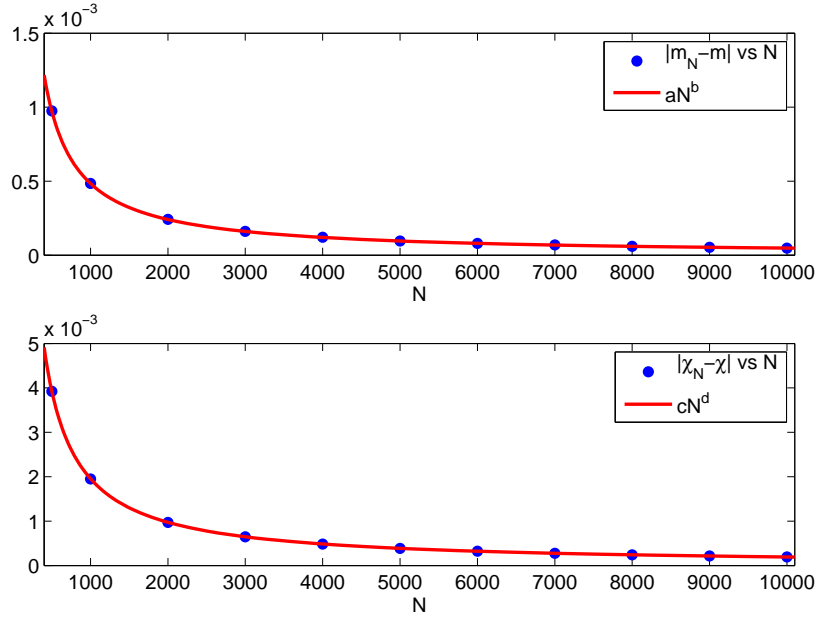
Figure 2: $J = 1.2$ and $h = 0.3$. Upper panel: $|m_N - m|$ as a function of N together with the best fit aN^b for the data in the right upper panel of fig.1. We obtain $a = 0.5047 \pm 0.0037$ and $b = -1.006 \pm 0.002$ with a goodness of fit $R^2 = 1$. Lower panel: $|\chi_N - \chi|$ as a function of N together with the best fit cN^d for the data in the right lower panel of fig.1. We obtain $c = 2.037 \pm 0.019$ and $d = -1.006 \pm 0.002$ with a goodness of fit $R^2 = 1$.

Gibbs probability distribution function $P_{N,J,h}$. Given (J, h), we consider 20 M-sample and we solve the maximum likelihood model for each M-sample independently; then we average the inferred values J_{exp} and h_{exp} of the model parameters, given by (12) and (13), over the 20 M-samples. We consider $J \in [0.6, 1.2]$ and $h \in [-0.3, 0.3]$. The obtained values for the case $h = 0.1$ and $h = -0.1$ are shown in fig.4 and in fig.5, where J_{exp} and h_{exp} are plotted as functions of J. Note that the inferred values of the parameters are in optimal agreement with the exact values (continuous lines in fig4 and in fig.5). The J_{exp} and h_{exp} values generated for $h = 0.3$ are reported in fig.6 to show that when the external field is enforced the prediction is quite good too: the points lie on the line of the exact agreement (red continuous lines in fig.6) even if the error bars are bigger than in previous cases.
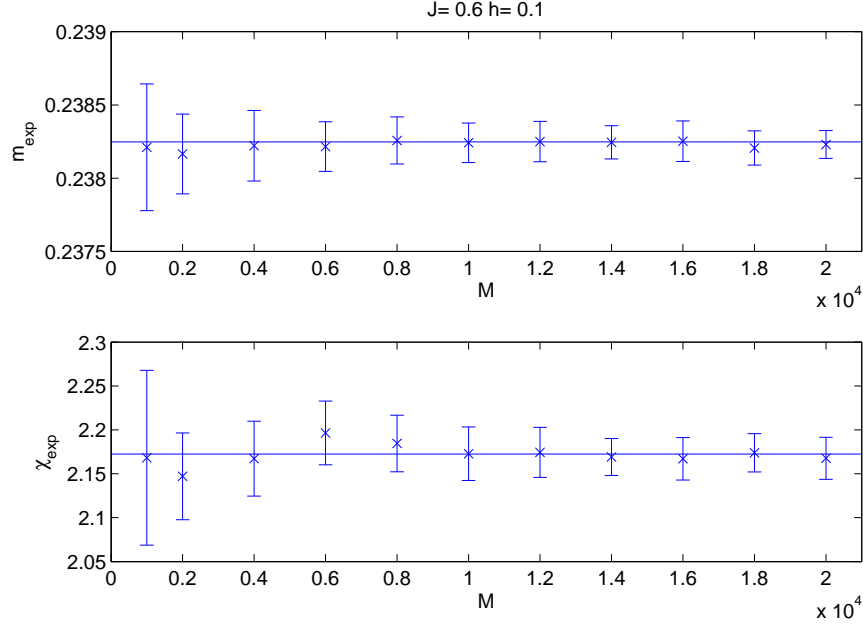
Figure 3: $N = 10000$, $J = 0.6$ and $h = 0.1$. Upper panel: Average magnetization m_{exp} (blue crosses) as a function of M (number of the configurations in the sample) together with statistical error bars over 20 M-sample for the Curie-Weiss model. The blue continuous line represents the finite size magnetization m_N for $N = 10000$. Lower panel: Susceptibility χ_{exp} (blue crosses) as a function of M (number of the configurations in the sample) together with statistical error bars over 20 M-sample. The blue continuous line represents the finite size susceptibility χ_N for $N = 10000$
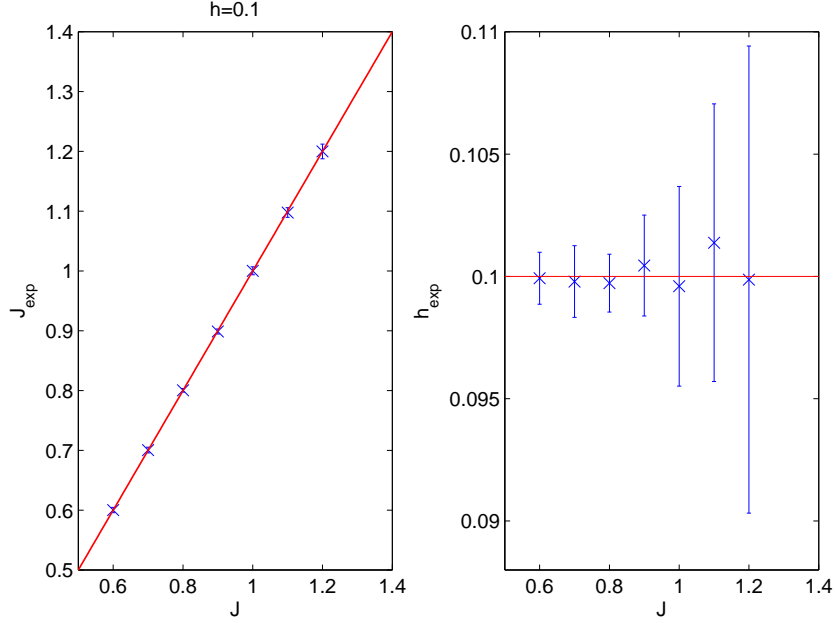
Figure 4: Left panel: J_{exp} as a function of J for $h = 0.1$ (blue crosses). Error bars are standard deviations on 20 different M-samples of the same simulation (see text for the details of the simulation). The red continous line represents $J_{exp} = J$. Right panel: The value of h_{exp} (blue crosses) calculated from (8) for the values of J_{exp} in the left panel, as a function of J together with the statistical error over 20 M-samples. The horizontal line corresponds to the exact value of the magnetic field $h = 0.1$

As a test problem for the multi-species mean-field model we consider a system of $N = 2000$ particles divided into $k = 2$ equally populated subsets ($N_1 = N_2 = 1000$) and a sample of $M = 10000$ independent spin configurations. Starting from 20 different couples of given values for the reduced interaction matrix

$$\mathbf{J} = \begin{pmatrix} J_{11} & J_{12} \\ J_{12} & J_{22} \end{pmatrix} \qquad J_{11}, J_{22} \in [0.55, 1.2], \quad J_{12} \in [-0.6, 1.1] \tag{27}$$

and for the external vector field

$$\mathbf{h} = \begin{pmatrix} h_1 \\ h_2 \end{pmatrix} \qquad h_1, h_2 \in [-0.3, 0.3] \tag{28}$$

we consider 20 M-samples for each couple $(\mathbf{J}, \mathbf{h})$ and we solve the maximum likelihood model for each one of them independently; then we average the inferred values $\mathbf{J}_{exp}$ and $\mathbf{h}_{exp}$ of the model parameters, given by (25) and (26), over the 20 M-samples (as in the one population model). In fig.7 the euclidean distances between $\mathbf{J}_{exp}$ and the initial reduced interaction matrix $\mathbf{J}$ (blue stars) and between $\mathbf{h}_{exp}$ and the initial external vector field $\mathbf{h}$ (red circles) are shown for each of the 20 choices of $\mathbf{J}$ and $\mathbf{h}$ (cases). We observe that, as in the one-population model (Curie-Weiss model), the inverse problem procedure to infer the experimental values for the coupling matrix and for the external field produces results in very good agreement with the initial values. In order to have a quantitative measure of the goodness of this procedure, we focus on the two most representative cases.

If we consider the case 1, for which the distance between the initial data matrix $\mathbf{J}$ and the inferred matrix $\mathbf{J}_{exp}$ gives the maximum value (see absolute errors in fig.7), we have

$$\mathbf{J} = \begin{pmatrix} 1.2 & 0.98 \\ 0.98 & 0.8 \end{pmatrix} \qquad \mathbf{J}_{exp} = \begin{pmatrix} 1.173 \pm 0.036 & 0.993 \pm 0.028 \\ 0.993 \pm 0.028 & 0.794 \pm 0.040 \end{pmatrix} \tag{29}$$

and for the corresponding values of the external field

$$\mathbf{h} = \begin{pmatrix} 0.1 \\ 0.2 \end{pmatrix} \qquad \mathbf{h}_{exp} = \begin{pmatrix} 0.102 \pm 0.012 \\ 0.198 \pm 0.011 \end{pmatrix}. \tag{30}$$

The errors on each value of the matrix $\mathbf{J}_{exp}$ and the vector $\mathbf{h}_{exp}$ are standard deviations across 20 different M-sample of the same $(\mathbf{J}, \mathbf{h})$-simulation.
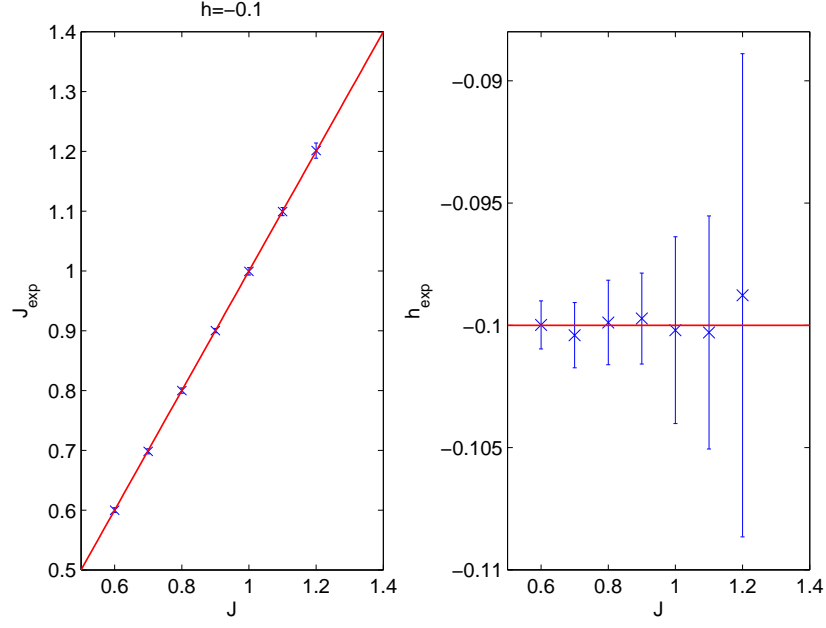
Figure 5: Left panel: J_{exp} as a function of J for $h = -0.1$ (blue crosses). Error bars are standard deviations on 20 different M-samples of the same simulation (see text for the details of the simulation). The red continuous line represents $J_{exp} = J$. Right panel: The value of h_{exp} (blue crosses) calculated from (8) for the values of J_{exp} in the left panel, as a function of J together with the statistical error over 20 M-samples. The horizontal line corresponds to the exact value of the magnetic field $h = -0.1$
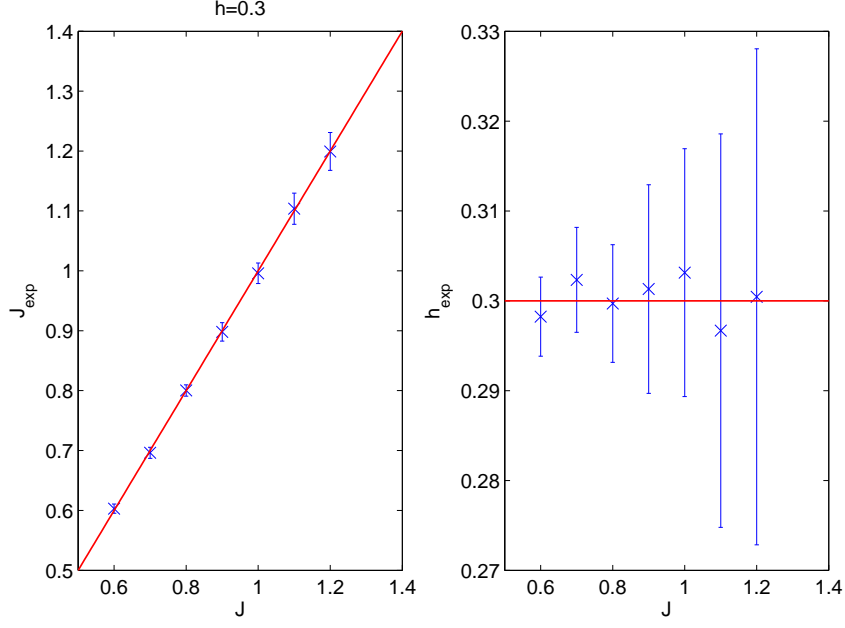
Figure 6: Left panel: J_{exp} as a function of J for $h = 0.3$ (blue crosses). Error bars are standard deviations on 20 different M-samples of the same simulation (see text for the details of the simulation). The red continuous line represents $J_{exp} = J$. Right panel: The value of h_{exp} (blue crosses) calculated from (8) for the values of J_{exp} in the left panel, as a function of J together with the statistical error over 20 M-samples. The horizontal line corresponds to the exact value of the magnetic field $h = 0.3$
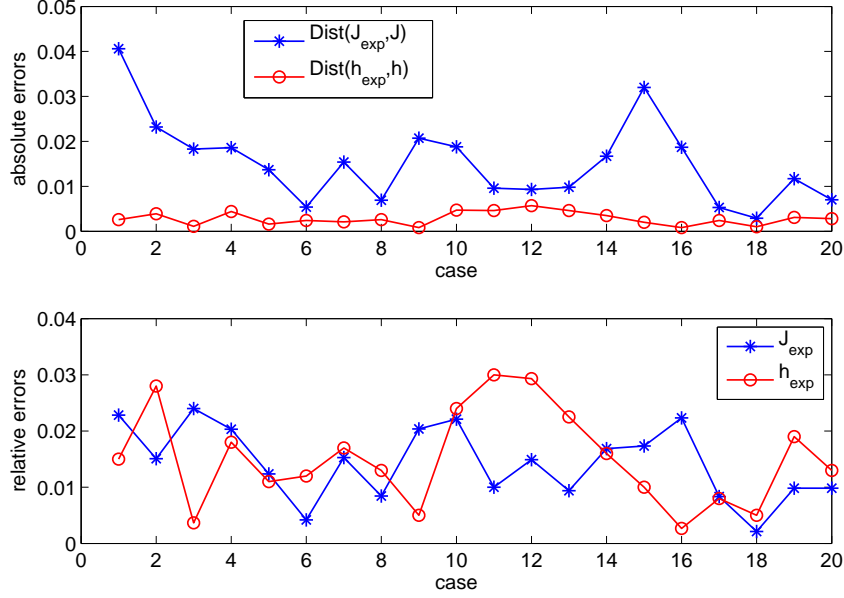
Figure 7: Upper panel: Absolute errors in reconstructing $\mathbf{J}$ and $\mathbf{h}$. Distance between the reconstructed matrix $\mathbf{J}_{exp}$ and the initial data matrix $\mathbf{J}$ (blue stars) and distance between $\mathbf{h}_{exp}$ and the initial $\mathbf{h}$ (red circle) for 20 different choices of parameters $\mathbf{J}$ and $\mathbf{h}$ (cases). The values of $\mathbf{J}_{exp}$ and $\mathbf{h}_{exp}$ are averaged across 20 M-sample (see Section 4 in the text for the details of the simulations). Lower panel: Relative errors in reconstructing $\mathbf{J}$ and $\mathbf{h}$. Maximum percentage errors for the reconstructed matrix $\mathbf{J}_{exp}$ (blue stars) and vector $\mathbf{h}_{exp}$ (red circle) for the same 20 multi-species cases considered in the upper panel.

The case 18 in fig.7, which gives the minimum value for the distance between $\mathbf{J}$ and $\mathbf{J}_{exp}$ corresponds to

$$\mathbf{J} = \begin{pmatrix} 0.6 & -0.8 \\ -0.8 & 0.9 \end{pmatrix} \qquad \mathbf{J}_{exp} = \begin{pmatrix} 0.601 \pm 0.022 & -0.798 \pm 0.019 \\ -0.798 \pm 0.019 & 0.901 \pm 0.020 \end{pmatrix} \tag{31}$$

and to the external field

$$\mathbf{h} = \begin{pmatrix} -0.2 \\ -0.3 \end{pmatrix} \qquad \mathbf{h}_{exp} = \begin{pmatrix} -0.201 \pm 0.005 \\ -0.300 \pm 0.005 \end{pmatrix}. \tag{32}$$

The errors on each value of the matrix $\mathbf{J}_{exp}$ and $\mathbf{h}_{exp}$ are the standard deviations across 20 different M-samples of the same $(\mathbf{J}, \mathbf{h})$-simulation.

5   Conclusions and perspectives

In this paper we have tested the robustness of the inversion method in a class of statistical mechanics mean field models. The novelty of the results is both on the finite size behavior of the exact solutions and on the quality of the inversion for finite number of samples. Our findings show that with a modest investment on samples we are able to reconstruct the values of the parameters within a few percentages. The relevance of the problem comes from the necessity to have a fully tested method in the parameter evaluation from real data of socio-economic type, as started from the seminal work of Brock and Durlauf. The nature of the investigated model belongs to those without intrinsic randomness but we plan to extend a similar analysis to those cases with random interactions like the Sherrington-Kirkpatrick model, and/or random network connections among agents, like in the diluted models.

References

[1] W. Bialek, A. Cavagna, I. Giardina, T. Mora, E. Silvestri, M. Viale and A.M. Walczak, Statistical mechanics for natural flocks of birds, Proceedings of the National Academy of Sciences, **109**, 4786-4791, (2012)

[2] W.A. Brock and S.N. Durlauf, Discrete choice with social interactions, Rev. Economic Studies, **68**, 235260, (2001)

[3] P. Contucci and S. Ghirlanda, Modeling Society with Statistical Mechanics: an Application to Cultural Contact and Immigration, Quality and Quantity, **41**, 569-578, (2007)

[4] R.S. Ellis, Entropy, large deviations, and statistical mechanics, Springer (2005)

[5] R.S. Ellis, C.M. Newman and J.S. Rosen, Limit theorems for sums of dependent random variables occurring in statistical mechanics, Probability Theory and Related Fields, **51**, 153-169, (1980)

[6] M. Fedele, A mean field model for the collective behaviour of interacting multispecies particles: mathematical results and application to the inverse problem, PhD thesis, University of Bologna, (2011)

[7] R.A. Fisher, Theory of statistical estimation, Mathematical Proceedings of the Cambridge Philosophical Society, **22**, 700–725, (1925)

[8] I. Gallo and P. Contucci, Bipartite mean field spin systems. Existence and solution, Mathematical Physics Electronic Journal, **14**, 1-22, (2008)

[9] R. B. Griffiths, A proof that the free energy of a spin system is extensive, J. Math. Phys, **5**, 1215-1222, (1964)

[10] R. B. Griffiths, Correlation in Ising Ferromagnets, J. Math. Phys, **8**, 478-483, (1967)

[11] F.E.Harrel, Regression modeling strategies: with applications to linear models, logistic regression and survival analysis, Springer Series in Statistics, 2001

[12] D.G.Kelly and S. Sherman, General Griffiths' Inequalities on Correlations in Ising Ferromagnets, J. Math. Phys. **9**, 466, (1968)

[13] E.T. Janes, Information theory and statistical mechanics, Physical review, **106**, 620, (1957)

[14] S. Loreti, Il problema inverso in meccanica statistica, Laurea Thesis, University of Bologna, (2010)

[15] D. McFadden, Economic Choices, American Economic Review, **91**, 351-378 (2001)

[16] Q. Michard and J.-P. Bouchaud, Theory of collective opinion shifts: from smooth trends to abrupt swings, European Physical Journal B **47**, 151-159, (2005).

[17] R. Monasson and S. Cocco, Fast inference of interactions in assemblies of stochastic integrate-and-fire neurons from spike recordings, Journal of computational neuroscience, **31**, 199-227, (2011)

[18] Y. Roudi, J. Tyrcha and J. Hertz, Ising model for neural data: Model quality and approximate methods for extracting functional connectivity, Physical Review E, **79**, 051915, (2009)

[19] V. Sessak and R. Monasson, Small-correlation expansions for the inverse Ising problem, Journal of Physics A: Mathematical and Theoretical, **42**, 055001, (2009)

[20] S. Galam, Y. Gefen and Y. Shapir, A mean behavior model for the process of strike, Mathematical Journal of Sociology **9**, 1-13, (1982)